Surface weighted average concentration (SWAC): New jargon for an old idea

John W. Kern,, Kern Statistical Services, Inc. Sauk Rapids, MN, USA

Richard E. Remington, Quantified, Inc. Boise, ID, USA Todd W. King, CDM, Detroit, MI, USA

Introduction

The United States Environmental Protection Agency (USEPA 1992) coined the term "exposure point concentration" (EPC) to describe the interaction between the spatial distribution of contaminants and organism use. The EPC is an upper confidence limit for the arithmetic mean intended to represent exposure for risk assessment.

Increasingly, at large contaminated sediment sites such as the Fox and Kalamazoo Rivers, EPCs have been estimated using what is termed surface weighted average concentration (SWAC). Reible et al. (2003) described SWAC as a useful surrogate risk metric, representing the average contaminant concentration in the biologically active portion of sediment. In this sense SWAC is used in place of an EPC for quantifying exposure, yet SWAC is routinely applied without measures of uncertainty. In essence, the EPC approach provides a margin of safety (upper confidence limit), while SWAC approaches typically ignore uncertainty in favor of the perceived precision of interpolated concentrations.

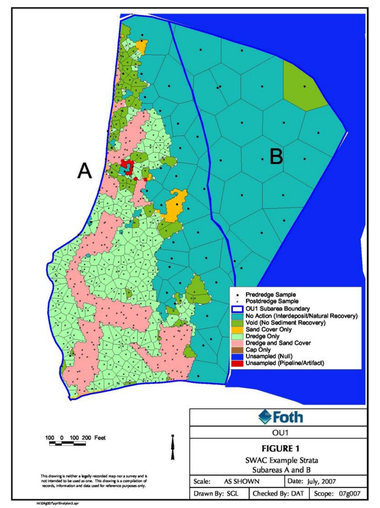


Fig. 1. Target SWAC levels are part of the remedial action objectives at the Fox River in Wisconsin. Thiessen polygons have been proposed as a basis for weighted averages in OU1. *Included with permission from Foth (2007).*

Practical Problems

- SWAC jargon has not been consistently applied.
  - Sometimes used interchangeably with EPC but without accounting for uncertainty
  - At other times used to specify a particular calculation procedure
- Many calculation procedures have been applied.
  - Average of an interpolated surfaces (IDW, Kriged, Natural Neighbor)
  - Weighted average based on Thiessen Polygons
  - Geostatistical procedures (Block Kriging, Conditional Simulation)
- Precision is rarely quantified
  - Data objectives have been difficult to establish
  - Geo-spatial calculations are often perceived to be preferred
  - Attempts to accommodate biased sample data often drives method selection
- Ill-defined arguments over "*correct*" procedures slow progress.
  - Arguments persist over differences that often are smaller than the limits of precision
- Importance of sampling design is often overlooked or ignored

Objectives

- Demonstrate use of conditional simulation for estimating the number of samples necessary to achieve a specified precision for non-standard sampling and estimation methods.
- Compare precision of SWAC estimators—arithmetic average, Thiessen polygons and natural neighbor interpolation
- Evaluate performance of SWAC estimators for simple random, and systematic sampling designs (SRS and SYS respectively)
- Evaluate performance of SWAC estimators with inclusion of a 20% biased sub-sample with SRS and SYS designs

Methods

- Sample data from the Plainwell Impoundment, Kalamazoo River, Michigan was used to develop a geostatistical probability model describing the spatial distribution of PCB concentrations in the top 6 inches of sediment in formerly impounded sediments
- Geographic coordinates were transformed to long- and cross-flow coordinates prior to analysis to improve consistency between geostatistical model and geomorphic processes

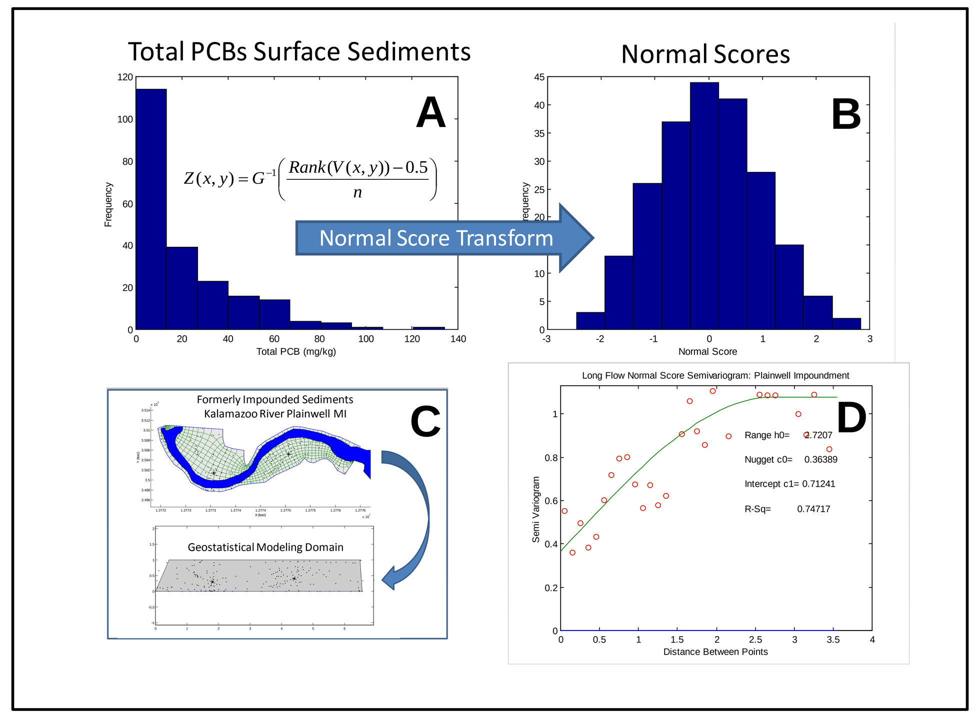
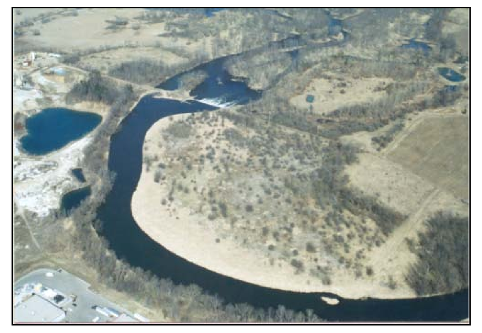


Fig. 2. Surface total PCB concentrations (Panel A) were normal score transformed (Panel B). Geographic coordinates were "straightened" (Panel C) and the semivariogram for the transformed data was estimated (Panel D).

- Multiple equally likely maps of PCB concentration were simulated from this model using the p-field simulation
- Maps were sampled using SRS and SYS designs and SWAC estimators were applied to each sample



- SWAC estimates were compared with actual averages based on the simulated map providing estimates of bias and precision for each method
- Simulations were repeated for a range of sample sizes

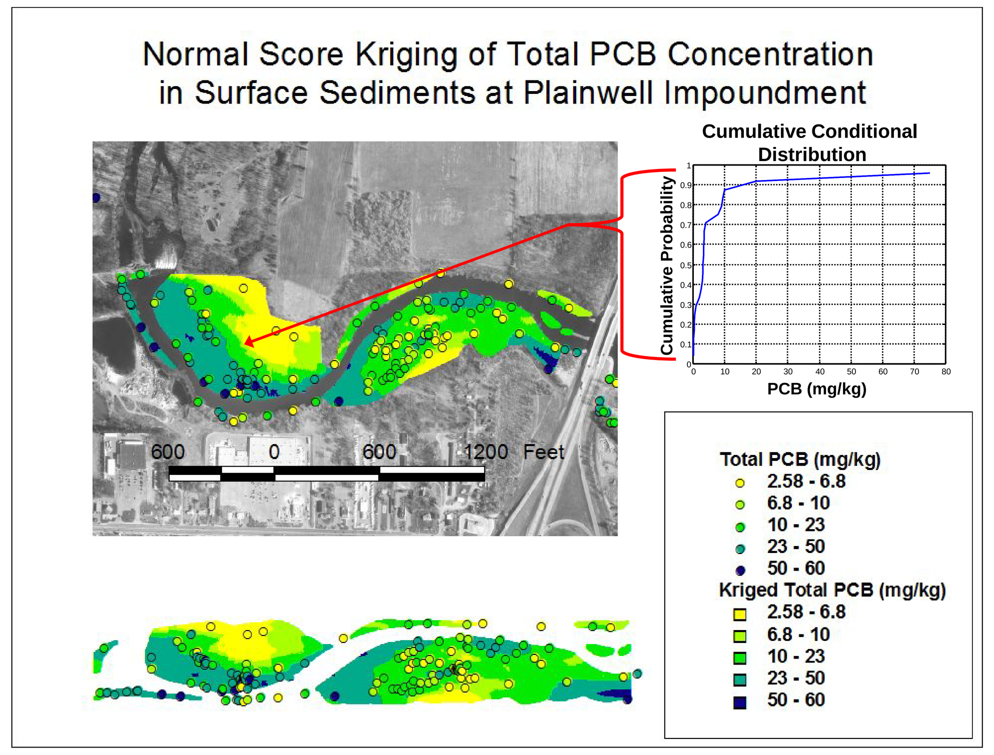


Fig. 3. Kriged surface based on geostatistical multi-Gaussian model.

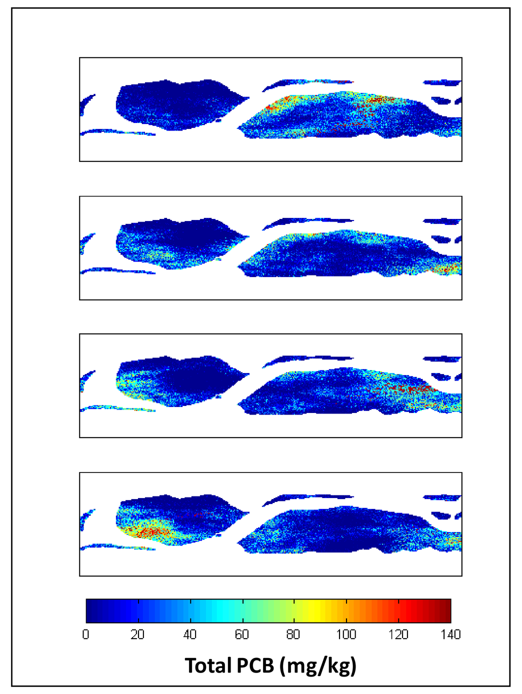


Fig. 4. Four equally likely maps of total PCB concentration drawn from the multi-Gaussian model.

Conditional Simulation

- P-field simulation with the Fast Fourier Transform method was used to sample from the cumulative distribution functions estimated from normal score Kriging.
- 1000 equally likely maps were simulated consistent with data
  - Histogram
  - Semivariogram
  - Interpolation of sample data
- Sixty combinations of estimation method, sampling design and sample size were calculated
- True SWAC value was estimated from the exhaustive map
- Bias and standard deviation of estimated SWAC was calculated

Results

Estimation Bias

- Arithmetic average and Thiessen Polygons were unbiased for unbiased designs
- Natural neighbor method was biased slightly high for the SYS design
- The arithmetic average overestimated the true mean for biased designs
- The Thiessen polygon and natural neighbor methods underestimated the true mean for the SRS design with biased subsample
- For the SYS design with biased subsample the Thiessen polygon method underestimated the true mean, while the natural neighbor method was essentially unbiased

Precision (Relative Error)

- Sample size was the primary driver of precision
- Estimators based on biased designs were more variable (i.e. higher relative error)
- Estimators based on SYS were less variable than those based on SRS
- For SRS designs, relative error was similar for all three estimators
- Relative error for the natural neighbor estimator was substantially higher than the arithmetic average and Thiessen polygon methods with SYS designs

Sample Size Determination

- 40-50 sample locations were adequate to achieve 15% relative error.

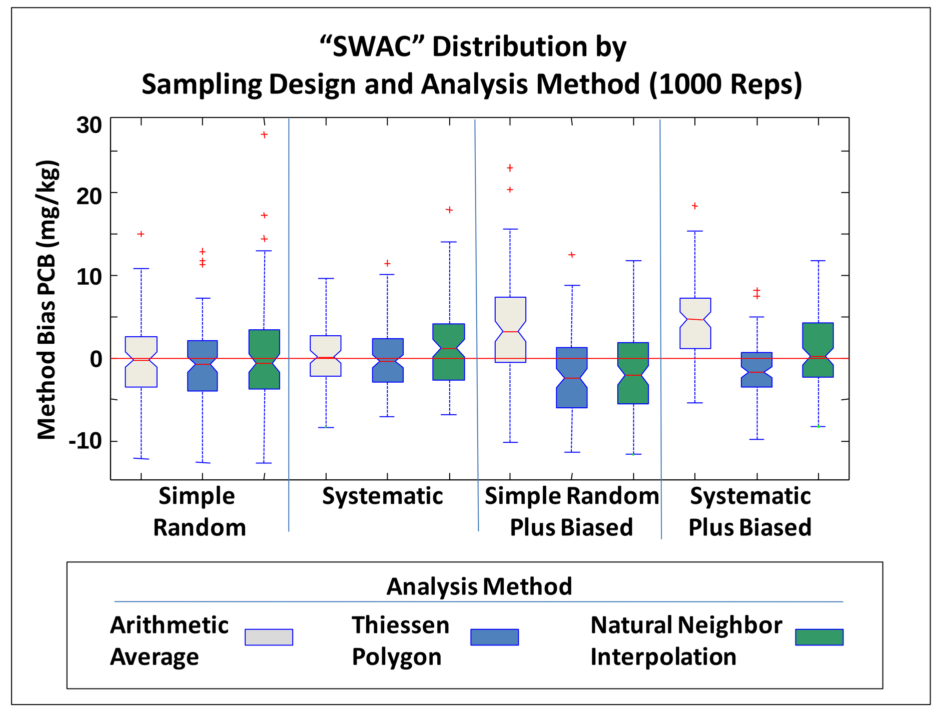


Fig. 5. Estimation bias for arithmetic average, Thiessen polygon and natural neighbor estimators with biased and unbiased simple random and systematic sampling designs.

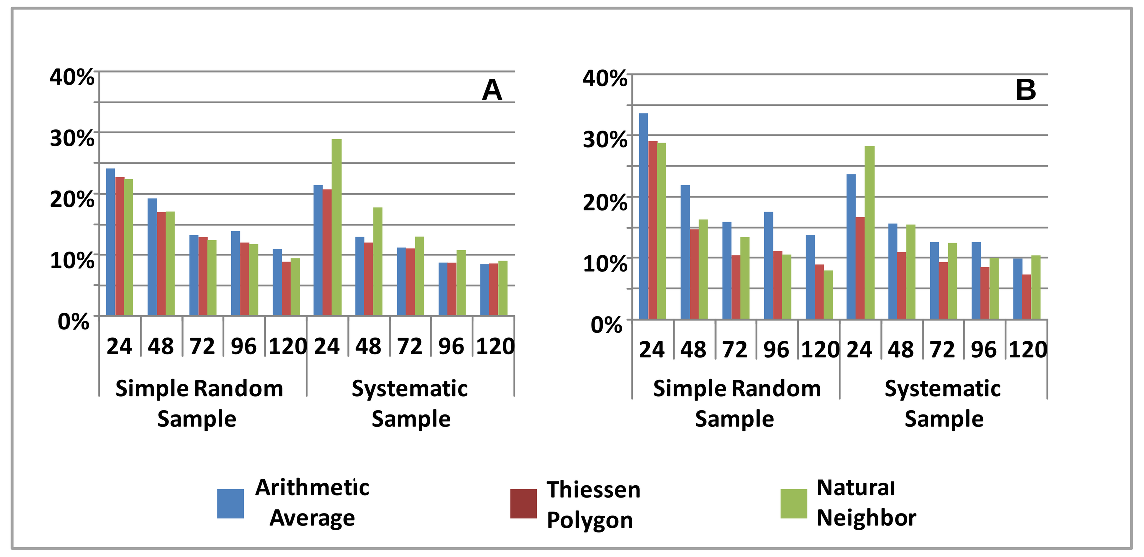


Fig. 6. Relative error (Standard Error/SWAC) for arithmetic average, Thiessen polygon and natural neighbor estimators for unbiased (Panel A) and biased (Panel B) sampling designs.

Discussion

This study illustrates that geostatistical simulation can be used to identify important properties of candidate estimators and to select appropriate sample numbers to meet specified data quality objectives (DQOs).

Efforts to estimate SWAC at large contaminated sites often focus on methods when sample size and design may be the more important factors limiting accuracy and precision.

Estimating trends in surface concentrations of contaminants is of increasing interest for demonstrating remedial effectiveness. To make statistical comparisons, estimates of uncertainty in SWAC are necessary. The estimators discussed in this study and most other SWAC estimators such as block Kriging are linear estimators of the form

$$G(x,y)=\sum_{i=1}^{n} w_i f(x_i,y_i)$$

and therefore sampling variance and confidence intervals can be estimated by standard formulas developed for global block Kriging or using conditional simulation. Kern (2000) illustrates estimation of confidence limits and Kern (2001) illustrates incorporation of SWAC uncertainty into remedial alternative selection.

Use of biased sampling designs inflated estimation variance and caused biases that were not completely corrected by the Thiessen or natural neighbor methods. To avoid these problems, SWAC estimation at larger scales should be based on unbiased sampling designs.

Methods and results discussed here can be extended to estimation of volume by substituting depth of contamination (DOC) for concentration. This suggests that estimation of total in situ volume of contaminated sediment may be best approached by estimating global mean DOC using linear estimators, rather than using complex 3-D models. An advantage to this approach is that confidence limits are readily available using standard geostatistical procedures.

These results are conditional on the statistical distribution of PCB data at Plainwell Impoundment. However, the PCB distribution is typical of that found at many other large PCB contaminated sites. This approach could be replicated at other large sediment sites to develop general guidance for planning sediment investigations intended to estimate SWAC or for remedy selection.

Literature cited

Reible, D.D., et al. (2003). Comparison of long –term risks of removal and in-situ management of contaminated sediments in the Fox River. *J. Soil Contaminat*, 12(3), 325-344.

Foth Infrastructure and Environment, 2007. 2007 Design Supplement, Appendix C, OU1 SWAC White Paper.

Kern, J. W. And K. O. Coyle, 2000. Estimation of the global block Kriging mean with large sample sizes on irregular polygonal regions: Applications to acoustic surveys for zooplankton in the Western Aleutian Islands. *Canadian Journal of Fisheries and Aquatic Sciences.* Vol 57: 2112-2121.

Kern, J. W. et al. (2001). Integrating statistical uncertainty into cost benefit analysis of remedial alternatives. Proceedings of the 22nd annual meeting of the Society of Environmental Toxicology and Chemistry. Baltimore MD.

USEPA 1992. Supplemental guidance to RAGS: Calculating the concentration term.   Intermittent Bulletin Volume 1, Number 1. Publication: 9285.7-081, May 1992.

Contact: jkern@KernStat.com.

Download PDF-version: www.KernStat.com/Download/SWAC2009.pdf